DOCUMENT RESUME

ED 217 083 TM 820 408

AUTHOR Gilmer, Jerry S.; Feldt, Leonard S.
TITLE The Standard Errors of the Feldt-Gilmer Congeneric Reliability Coefficients: Iowa Testing Programs Occasional Papers. Number 31.
INSTITUTION Iowa Testing Programs, Iowa City.
PUB DATE Feb 82
NOTE 44p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Educational Testing; *Error of Measurement; *Mathematical Formulas; Mathematical Models; *Measurement Techniques; *Statistical Analysis; *Test Reliability
IDENTIFIERS *Congeneric Tests; Monte Carlo Methods; Reliability Formulas; *Tau Equivalence; Test Length

ABSTRACT

The Feldt-Gilmer congeneric reliability coefficients make it possible to estimate the reliability of a test composed of parts of unequal, unknown length. The approximate standard errors of the Feldt-Gilmer coefficients are derived via a method using the multivariate Taylor's expansion. Monte Carlo simulation is employed to corroborate the theoretical approximations for eight hypothetical tests. The Feldt-Gilmer coefficients are shown to be appropriate with both congeneric and tau-equivalent parts. Their superiority over the special case of Kristof's coefficient increases with the number of parts. Their standard errors also compare favorably to that of Cronbach's alpha coefficient when applied to tests with tau-equivalent parts. Tables illustrate the charts. (Author/CM)

ED217083

IOWA TESTING PROGRAMS OCCASIONAL PAPERS

Number 31 - February 1982

The Standard Errors of the Feldt-Gilmer Congeneric Reliability Coefficients

Jerry S. Gilmer

University of Arizona

Leonard S. Feldt

University of Iowa

Abstract

The Feldt-Gilmer congeneric reliability coefficients make it possible to estimate the reliability of a test composed of parts of unequal, unknown length. These coefficients include Kristof's coefficient and Cronbach's alpha coefficient as special cases. In this paper the approximate standard errors of the Feldt-Gilmer coefficients are derived via a method that uses the multivariate Taylor's expansion. Monte Carlo simulation is employed to corroborate the theoretical approximations for eight hypothetical tests. It is shown that the Feldt-Gilmer coefficients are appropriate with both congeneric and tau-equivalent parts. Their superiority over the Kristof coefficient increases with the number of parts ($K \geq 4$). Their standard errors also compare favorably to that of Cronbach's alpha when applied to tests with tau-equivalent parts.

Introduction

Measurement situations sometimes occur in which a multiple-part measure can be assumed to consist of *congeneric* parts, but not *essentially tau-equivalent* parts. Congeneric parts are parallel in content and measure the same attributes, but differ in length. Tau-equivalent parts are parallel in content and functionally equal in length. More formally, parts j and h are congeneric if their true scores satisfy the relation $T_j = aT_h + b$, where a is a constant not necessarily equal to 1.0 and b is a constant not necessarily equal to 0 . Parts j and h are tau-equivalent if $a = 1.0$. For both types of parts error score variances and observed score variances may be heterogeneous from part to part. For congeneric parts the true score variances are, in general, heterogeneous. For tau-equivalent parts true score variances are homogeneous.

Examples of congeneric parts are not difficult to find. In reading tests the passages often vary in length and in the number of items based on each passage. If these numbers of items vary appreciably, the passage scores will very likely be congeneric. The ratings of individual judges on a panel will be congeneric if some judges concentrate their ratings in the midd'- of the scale and avoid extreme ratings, while others spread their ratings over the full range. In an essay test, if the several questions differ in difficulty or are given different point values by the examiner, the scores on the various questions might well be congeneric rather than essentially tau-equivalent.

Coefficient alpha [Cronbach, 1951] provides an acceptable estimate of reliability if the part scores are essentially tau-equivalent [Novick and Lewis, 1967]. Horst's coefficient [1951] will serve for congeneric parts if the number of items can be assumed to represent the effective or functional length of each part. In the case of ratings and essay tests, however, the lengths of the parts cannot be inferred from any obvious feature of the parts. In such situations, one must use one or another of the approaches developed by Feldt [1975], Kristof [1974], or Gilmer and Feldt [1981].

The first purpose of this study was to develop approximate standard error formulas for the last two of these congeneric reliability coefficients. The basic technique used for this purpose is the "delta" method of Kendall and Stuart [1969, Volume I, Chapter 10]. This method makes it possible to approximate the standard error of certain statistics (to order N^{-1}) even though the sampling distribution of the statistic is unknown. The second purpose was to verify the comparative values of these standard errors through a Monte Carlo simulation of the sampling distributions of the coefficients. The last purpose was to draw some conclusions regarding the choice of estimates, on the basis of their bias and standard error.

The Congeneric Coefficients

Because the derivations of these coefficients have been presented in detail elsewhere [Gilmer, 1981], they will not be repeated here.

However, the notational scheme and the structure of the congeneric model must be made explicit.

The congeneric model can be specified as follows for a test that consists of K parts:

$$X = X_1 + X_2 + X_3 + \cdots + X_K$$

$$T = T_1 + T_2 + T_3 + \cdots + T_K$$

$$E = E_1 + E_2 + E_3 + \cdots + E_K$$

$$X = T + E$$

$$X_j = T_j + E_j = \lambda_j T + b_j + E_j \qquad (j = 1, \ldots, K)$$

$$\Sigma\lambda_j = 1.0, \; \lambda_j > 0$$

$$\Sigma b_j = 0$$

In this representation X, T, and E are the observed, true, and error scores for the total test, respectively; X_j , T_j , and E_j $(j = 1, \ldots K)$ are part-test observed, true, and error scores. The λ_j are constants which represent the proportions of the total test true score that are contributed by the various parts. They may be viewed as the "proportionate functional lengths" of the parts. The b_j are constants which make allowance for the differences in mean score on the parts, beyond differences arising from variation in part-test length. The error scores, E_j , are assumed to be mutually independent and independent of true scores. The linear relationship among the part-test true scores implies that they are perfectly correlated with the true scores on the full test.

Additional parameters of importance are as follows:

σ_{ij} = covariance of parts i and j

σ_{jj} = variance of part j

σ_T^2 = total test true score variance, and

σ_X^2 = total test observed score variance.

Given the independence of E_i , E_j , and T_j the observed score variances and covariances are as follows:

$$\sigma_{jj} = \lambda_j^2\sigma_T^2 + \sigma_{E_j}^2 \qquad (j = 1, \ldots, K),$$

$$\sigma_{ij} = \lambda_i \lambda_j \sigma_T^2 \qquad (i \neq j)$$

The total of the covariances in row i of the part-test covariance matrix is

$$\sum_{\substack{j \\ i \neq j}} \sigma_{ij} = \lambda_i \Sigma\lambda_j \sigma_T^2 = \lambda_i(1-\lambda_i)\sigma_T^2$$

In the part-test covariance matrix for any examinee sample, let the row with the largest sum of covariances be designated as row ℓ . Designate by C_j the quotient obtained by dividing the sum of covariances in row j by the sum of covariances in row ℓ . Let $\hat{Y} = \hat{\lambda}_\ell(1-\hat{\lambda}_\ell)$. Then,

$$C_1 = \sum_{j \neq 1} \hat{\sigma}_{1j} \Big/ \sum_{j \neq \ell} \hat{\sigma}_{\ell j} = \hat{\lambda}_1(1-\hat{\lambda}_1) \Big/ \hat{\lambda}_\ell(1-\hat{\lambda}_\ell) = \hat{\lambda}_1(1-\hat{\lambda}_1) \Big/ \hat{Y}$$

$$\ldots \qquad \ldots \qquad \ldots$$

$$C_\ell = \sum_{j \neq \ell} \hat{\sigma}_{\ell j} \Big/ \sum_{j \neq \ell} \hat{\sigma}_{\ell j} = 1.0$$

$$\ldots \qquad \ldots \qquad \ldots$$

$$C_K = \sum_{j \neq K} \hat{\sigma}_{Kj} \Big/ \sum_{j \neq \ell} \hat{\sigma}_{\ell j} = \hat{\lambda}_K(1-\hat{\lambda}_K) \Big/ \hat{\lambda}_\ell(1-\hat{\lambda}_\ell) = \hat{\lambda}_K(1-\hat{\lambda}_K) \Big/ \hat{Y}$$

Consider the pair of functions:

$$f_I(\hat{Y}) = \frac{K}{2} - 1 + \sqrt{.25-\hat{Y}} - \sum_{j\neq\ell} \sqrt{.25-C_j\hat{Y}} = 0 \text{ , and} \tag{1a}$$

$$f_{II}(\hat{Y}) = \frac{K}{2} - 1 - \sum_j \sqrt{.25-C_j\hat{Y}} = 0 \text{ .} \tag{1b}$$

One or the other of these functions must be solved for $\hat{Y}$, the choice being dictated by the algebraic sign of $f_I(.25)$. The solution for $\hat{Y}$ is then substituted into the following formula to obtain an estimate of the total test true score variance:

$$\hat{\sigma}_T^2 = \sum_{j\neq\ell} \hat{\sigma}_{\ell j} \Big/ \hat{Y} \text{ .} \tag{2}$$

Using the variance ratio definition of reliability, we then obtain the first congeneric coefficient:

$$r_{F1} = \sum_{j\neq\ell} \hat{\sigma}_{\ell j} \Big/ (\hat{Y})(\hat{\sigma}_X^2) \text{ .} \tag{3}$$

For the second congeneric coefficient, designated r_{F2} , we again identify the row with the largest sum of covariances as row ℓ . We then obtain the following quotients:

$$D_1 = \frac{\sum_{j\neq 1} \hat{\sigma}_{1j} - \hat{\sigma}_{1\ell}}{\sum_{j\neq\ell} \hat{\sigma}_{\ell j} - \hat{\sigma}_{1\ell}}$$

$$D_2 = \frac{\sum_{j\neq 2} \hat{\sigma}_{2j} - \hat{\sigma}_{2\ell}}{\sum_{j\neq\ell} \hat{\sigma}_{\ell j} - \hat{\sigma}_{2\ell}}$$

$$\cdots$$

$$D_\ell = \frac{\sum_{j\neq\ell} \hat{\sigma}_{\ell j}}{\sum_{j\neq\ell} \hat{\sigma}_{\ell j}} = 1.0$$

$$\cdots$$

$$D_K = \frac{\sum_{j\neq K} \hat{\sigma}_{Kj} - \hat{\sigma}_{K\ell}}{\sum_{j\neq\ell} \hat{\sigma}_{\ell j} - \hat{\sigma}_{K\ell}}$$

The second congeneric coefficient is then defined as

$$r_{F2} = \frac{(\Sigma D_j)^2}{(\Sigma D_j)^2 - \Sigma D_j^2} \cdot \frac{\hat{\sigma}_X^2 - \Sigma \hat{\sigma}_{jj}}{\hat{\sigma}_X^2} . \tag{4}$$

The last congeneric coefficient considered in this paper is that derived by Kristof [1974] for a three-part test. To apply this estimate of reliability to a longer test, one must combine parts and reduce the instrument to three parts. The formula for Kristof's coefficient is

$$r_K = \frac{(\hat{\sigma}_{12}\hat{\sigma}_{13} + \hat{\sigma}_{12}\hat{\sigma}_{23} + \hat{\sigma}_{13}\hat{\sigma}_{23})^2}{(\hat{\sigma}_{12})(\hat{\sigma}_{13})(\hat{\sigma}_{23})\hat{\sigma}_X^2} . \tag{5}$$

If a test consists of only three parts, sample estimates r_{F1}, r_{F2}, and r_K will be identical.

Approximate Sampling Variances

A method by which the sampling variance of a statistic can be determined when the sampling distribution is unknown is presented by Kendall and Stuart [1969, Volume I, Chapter 10]. This method, called the delta method, is applicable when the statistic of interest is a function of "simpler" statistics, the sample values of which are always greater than zero. With sufficiently large samples, this condition can be met in the present context.

The delta method uses the initial terms in the Taylor expansion of a function of several variables. Let $\underline{X} = (X_1, X_2, \ldots, X_p)'$ and $\underline{a} = (a_1, a_2, \ldots, a_p)'$. Then the terms through the first derivative are:

$$f(\underline{X}) = f(\underline{a}) + \sum_i^p (X_i - a_i) f_i'(\underline{a}) .$$

Subtracting $f(\underline{a})$ from both sides, we get

$$f(\underline{X})-f(\underline{a}) = \sum_{i}^{p}(X_i-a_i)f_i'(\underline{a})$$

and squaring both sides yields the relation

$$[f(\underline{X})-f(\underline{a})]^2 = \sum_{i}^{p}\sum_{j}^{p}(X_i-a_i)(X_j-a_j)f_i'(\underline{a})f_j'(\underline{a}) . \tag{6}$$

If $\underline{a} = \mathcal{E}(\underline{X})$ and $\mathcal{E}[f(\underline{X})] = f(\underline{a})$, then taking the expected value of both sides of (6) gives an expression for the variance of f(X) :

$$\begin{aligned} \mathrm{Var}[f(\underline{X})] &= \mathcal{E}[\sum_{i}\sum_{j}(X_i-a_i)(X_j-a_j)f_i'(\underline{a})f_j'(\underline{a})] \\ &= \sum_{i}\sum_{j}[\mathrm{Cov}(X_i,X_j)f_i'(\underline{a})f_j'(\underline{a})] . \end{aligned}$$

Define $\mathcal{E}(\underline{X}) = \underline{\theta}$ and $f(\underline{\theta}) = g$. We then have

$$\mathrm{Var}[f(\underline{X})] = \sum_{i}\sum_{j}\left[\frac{\partial g}{\partial\theta_i}\cdot\frac{\partial g}{\partial\theta_j}\cdot \mathrm{Cov}(X_i,X_j)\right] . \tag{7}$$

This last equation is the same as equation 10.12 of Kendall and Stuart [1969].

In the context of this paper the X_i represent the sample variances and covariances of the part tests, i.e., the elements in the sample variance-covariance matrix; the θ_i are the population variances and covariances of the part tests. The $f(\underline{X})$, in our case, is any one of the reliability estimates discussed previously. Using r as a particular reliability estimate and ρ as its population counterpart, we obtain from (7):

$$\mathrm{Var}(r) = \sum_{i}\sum_{j}\left[\frac{\partial\rho}{\partial\theta_i}\cdot\frac{\partial\rho}{\partial\theta_j}\cdot \mathrm{Cov}(X_i,X_j)\right] . \tag{8}$$

It is necessary, then, to determine $Cov(\tilde{X}_1, \tilde{X}_2)$, the covariance of two covariances. In particular, we need $Cov(S_{ij}, S_{mn})$, where $i,j,m,n, = 1, \cdots, K$; K = the number of part tests and S represents the sample estimate of σ_{ij}.

The determination of $Cov(S_{ij}, S_{mn})$ is not straightforward. Kendall and Stuart develop the covariance of two covariances with the use of Fisher's k-statistics [Fisher, 1928, in Kendall & Stuart, 1969, Volume I]. The end result of this extended derivation is as follows:

$$Cov(S_{ij}, S_{mn}) = \frac{\sigma_{im}\sigma_{jn} + \sigma_{in}\sigma_{jm}}{N-1} \qquad (9)$$

The variance obtained from (8) is not exact. It is correct to the order of N^{-1}. Therefore, the approximation and the true value of the sampling variance may differ by an amount that involves N^2, N^3, $\cdots$ in the denominator. With N relatively large, say $N \geq 100$, the difference would probably not be of great importance.

Theoretical Variance of r_{F1}

As in the foregoing section, we presume that the divisor for C_j is the maximum sum of row covariances. Without loss of generality we shall assume this maximum occurs for row 1. Since the delta method is applied to a function of parameters, the following relationships are defined among the parameters in the population variance-covariance matrix:

$$\alpha_1 = \sigma_{12} + \sigma_{13} + \cdots + \sigma_{1K} = \sum_{j\neq 1}\sigma_{1j}$$

$$\alpha_2 = \sigma_{21} + \sigma_{23} + \cdots + \sigma_{2K} = \sum_{j\neq 2}\sigma_{2j}$$

$$\cdots$$

$$\alpha_K = \sigma_{K1} + \sigma_{K2} + \cdots + \sigma_{K,K-1} = \sum_{j\neq K}\sigma_{Kj}$$

$$\gamma_1 = \frac{\alpha_1}{\alpha_1} = 1.0 ,$$

$$\gamma_2 = \frac{\alpha_2}{\alpha_1} ,$$

$$\gamma_3 = \frac{\alpha_3}{\alpha_1} ,$$

$$\cdots$$

$$\gamma_K = \frac{\alpha_K}{\alpha_1} .$$

The symbol γ_j is used here, rather than C_j, to emphasize that these are ratios of parameters. The population equation corresponding to equations (1a) and (1b) is

$$h(\sigma_{ij}) = g(\sigma_{ij},Y) = \frac{K}{2} - 1 \pm \sqrt{.25-Y} - \sum_{j\neq 1}\sqrt{.25-Y\gamma_j} = 0 . \tag{10}$$

The true score variance, σ_T^2, is α_1/Y, and the reliability is

$$\rho_{F1} = \frac{\sigma_T^2}{\sigma_X^2} = \frac{\alpha_1}{Y\sigma_X^2}$$

The variance of the sample estimate of ρ_{F1} is then given by

$$\mathrm{Var}(r_{F1}) = \sum_{b=1}^{KV}\sum_{d=1}^{KV}\left[\frac{\partial\rho_{F1}}{\partial\theta_b} \cdot \frac{\partial\rho_{F1}}{\partial\theta_d} \cdot \mathrm{Cov}(X_b,X_d)\right] , \tag{11}$$

In this expression $KV = K(K + 1)/2$, the total number of distinct elements in the variance-covariance matrix. The parameters θ_1 through θ_{KC} are the distinct part-test covariances ($KC = K(K-1)/2$); θ_{KC+1} through θ_{KV} are the part-test variances. The statistics X_1 through X_{KV} are the sample estimates corresponding to θ_1 through θ_{KV}. For example, if $K = 4$, the following would be defined:

$$
\begin{array}{ll}
KV = 10, & KC = 6, \\
\theta_1 = \sigma_{12} & X_1 = S_{12} \\
\theta_2 = \sigma_{13} & X_2 = S_{13} \\
\theta_3 = \sigma_{14} & X_3 = S_{14} \\
\theta_4 = \sigma_{23} & X_4 = S_{23} \\
\theta_5 = \sigma_{24} & X_5 = S_{24} \\
\theta_6 = \sigma_{34} & X_6 = S_{34} \\
\theta_7 = \sigma_{11} & X_7 = S_{11} \\
\theta_8 = \sigma_{22} & X_8 = S_{22} \\
\theta_9 = \sigma_{33} & X_9 = S_{33} \\
\theta_{10} = \sigma_{44} & X_{10} = S_{44}.
\end{array}
$$

Note that the first K-1 terms are the covariances from row 1, the covariances in the sum α_1. Note also that the elements in the lower triangle of the population variance-covariance matrix have not been included explicitly. This means that when we refer to θ_b ($b = 1, \cdots KV$), that is, when we refer to one of the σ_{ij}, we imply that $i \leq j$. This fact, that $i \leq j$, will appear again in this derivation.

The covariance of S_{ij} with S_{mn} is, from (9),

$$\mathrm{Cov}(S_{ij}, S_{mn}) = \frac{\sigma_{im}\sigma_{jn} + \sigma_{in}\sigma_{jm}}{N-1}$$

where N is sample size. Once $\partial\rho_{F1}/\partial\theta_b$ ($b = 1, \cdots, KV$) is found, the variance of r_{F1} can be approximately determined from expression (11) for any given population matrix and specified N .

The first step in this process is to apply the chain rule to $\partial\rho_{F1}/\partial\theta_b$:

$$\frac{\partial\rho_{F1}}{\partial\theta_b} = \frac{\partial \dfrac{\sigma_T^2}{\sigma_X^2}}{\partial\theta_b} = \frac{\sigma_X^2 \dfrac{\partial\sigma_T^2}{\partial\theta_b} - \sigma_T^2 \dfrac{\partial\sigma_X^2}{\partial\theta_b}}{(\sigma_X^2)^2} .$$

This result indicates that we need $\partial\sigma_X^2/\partial\theta_b$ and $\partial\sigma_T^2/\partial\theta_b$.

The first partial derivative is straightforward:

$$\frac{\partial\sigma_X^2}{\partial\theta_b} = \frac{\partial\left(\sum\limits_{i\neq j} \sigma_{ij} + \Sigma\sigma_{ii}\right)}{\partial\theta_b}$$

$$= \begin{cases} 2 \text{ if } b = 1, \cdots, KC \\ 1 \text{ if } b = KC + 1, \cdots, KV . \end{cases} \tag{12}$$

The partial derivative of $\partial\sigma_T^2/\partial\theta_b$ is not as direct:

$$\frac{\partial\sigma_T^2}{\partial\theta_b} = \begin{cases} \dfrac{\partial\sigma_T^2}{\partial\theta_b} , \text{ if } b = 1, \cdots, KC, \\ \\ 0 \quad , \text{ if } b = KC + 1, \cdots, KV . \end{cases} \tag{13}$$

We have then

$$\frac{\partial \rho_{F1}}{\partial \theta_b} = \begin{cases} \dfrac{\sigma_X^2 \dfrac{\partial \sigma_T^2}{\partial \theta_b} - 2\sigma_T^2}{\sigma_X^4}, & \text{if } b = 1, \cdots, KC \\[2em] \dfrac{-\sigma_T^2}{\sigma_X^4}, & \text{if } b = KC + 1, \cdots, KV. \end{cases} \tag{14}$$

We need to determine $\partial\sigma_T^2/\partial\theta_b$ for $b = 1, \cdots, KC$:

$$\frac{\partial \sigma_T^2}{\partial \theta_b} = \frac{\partial \frac{\alpha_1}{Y}}{\partial \theta_b} = \frac{Y \dfrac{\partial \alpha_1}{\partial \theta_b} - \alpha_1 \dfrac{\partial Y}{\partial \theta_b}}{Y^2}, \; b = 1, \cdots, KC. \tag{15}$$

We now need $\partial\alpha_1/\partial\theta_b$ and $\partial Y/\partial\theta_b$. The first of these is

$$\frac{\partial \alpha_1}{\partial \theta_b} = \frac{\partial (\sum'_{j \neq 1} \sigma_{1j})}{\partial \theta_b} \tag{16}$$

$$= \begin{cases} 1 \text{ if } b = 1, \cdots, K-1 \\ 0 \text{ if } b = K, \cdots, KC. \end{cases}$$

The determination of $\partial Y/\partial\theta_b$ involves implicit differentiation because Y is given implicitly in (10). We have from (10):

$$h(\sigma_{ij}) = g(\sigma_{ij}, Y) = 0.$$

Therefore,

$$\frac{dh}{d\theta_b} = \frac{dg}{d\theta_b} = \frac{\partial g}{\partial\sigma_{12}} \cdot \frac{\partial\sigma_{12}}{\partial\theta_b} + \frac{\partial g}{\partial\sigma_{13}} \cdot \frac{\partial\sigma_{13}}{\partial\theta_b} + \frac{\partial g}{\partial\sigma_{14}} \cdot \frac{\partial\sigma_{14}}{\partial\theta_b} + \cdots$$

$$+ \frac{\partial g}{\partial\sigma_{K-1,K}} \cdot \frac{\partial\sigma_{K-1,K}}{\partial\theta_b} + \frac{\partial g}{\partial Y} \cdot \frac{\partial Y}{\partial\theta_b} = 0,$$

which simplifies to

$$\frac{dh}{d\theta_b} = \frac{\partial g}{\partial\theta_b} + \frac{\partial g}{\partial Y} \cdot \frac{\partial Y}{\partial\theta_b} = 0, \quad b = 1, \cdots, KC.$$

Solving for $\partial Y/\partial\theta_b$, as required, we obtain:

$$\frac{\partial Y}{\partial\theta_b} = -\frac{\dfrac{\partial g}{\partial\theta_b}}{\dfrac{\partial g}{\partial Y}}. \tag{17}$$

Substituting the right sides of (16) and (17) into (15) we obtain:

$$\frac{\partial\sigma_T^2}{\partial\theta_b} = \frac{\partial\dfrac{\alpha_1}{Y}}{\partial\theta_b} = \frac{Y \cdot I + \dfrac{\dfrac{\partial g}{\partial\theta_b}}{\dfrac{\partial g}{\partial Y}} \cdot \alpha_1}{Y^2}$$

$$= \frac{\left(Y \cdot I \cdot \dfrac{\partial g}{\partial Y}\right) + \left(\dfrac{\partial g}{\partial\theta_b} \cdot \alpha_1\right)}{Y^2 \cdot \dfrac{\partial g}{\partial Y}}, \tag{18}$$

where

$I = 1$ if $b = 1, \cdots, K-1$

$I = 0$ if $b = K, \cdots, KC$.

Again, we need two partial derivatives: $\partial g/\partial\theta_b$, for $b = 1, \cdots, KC$, and $\partial g/\partial Y$.

Consider first $\partial g/\partial Y$. From (10) we can obtain:

$$\frac{\partial g}{\partial Y} = \pm \frac{\partial(\sqrt{.25-Y})}{\partial Y} - \sum_{k\neq 1} \frac{\partial(\sqrt{.25-Y\gamma_k})}{\partial Y}$$

where

$$\frac{\partial(\sqrt{.25-Y\gamma_k})}{\partial Y} = \frac{-\gamma_k}{2\sqrt{.25-Y\gamma_k}}$$

Thus, we have

$$\frac{\partial g}{\partial Y} = \mp \frac{1}{2\sqrt{.25-Y}} + \sum_{k\neq 1} \frac{\gamma_k}{2\sqrt{.25-Y\gamma_k}} \qquad (19)$$

The minus sign is used if $\lambda_1 > .5$, otherwise the plus sign is used.

Now consider the numerator of (17). From (10) we can obtain

$$\frac{\partial g}{\partial \theta_b} = -\sum_{k\neq 1} \frac{\partial(\sqrt{.25-Y\gamma_k})}{\partial \theta_b} ,$$

where

$$\frac{\partial(\sqrt{.25-Y\gamma_k})}{\partial \theta_b} = \frac{-Y}{\sqrt{.25-Y\gamma_k}} \cdot \frac{\partial \gamma_k}{\partial \theta_b} .$$

So we have

$$\frac{\partial g}{\partial \theta_b} = \sum_{k\neq 1} \frac{Y}{2\sqrt{.25-Y\gamma_k}} \cdot \frac{\partial \gamma_k}{\partial \theta_b} \qquad (20)$$

What now remains to be determined is $\partial\gamma_k/\partial\theta_b$. Applying the chain rule we obtain:

$$\frac{\partial\gamma_k}{\partial\theta_b} = \frac{\partial\frac{\alpha_k}{\alpha_1}}{\partial\theta_b} = \frac{\alpha_1\frac{\partial\alpha_k}{\partial\theta_b} - \alpha_k\frac{\partial\alpha_1}{\partial\theta_b}}{\alpha_1^2}, \; b = 1, \cdots, KC. \tag{21}$$

And once more we need two partial derivatives: $\partial\alpha_k/\partial\theta_b$ and $\partial\alpha_1/\partial\theta_b$, $b = 1, \cdots, KC$.

Consider $\partial\alpha_k/\partial\theta_b$ first:

$$\frac{\partial\alpha_k}{\partial\theta_b} = \frac{\partial(\sigma_{k1} + \sigma_{k2} + \cdots (-\sigma_{kk}) \cdots + \sigma_{kK})}{\partial\sigma_{ij}}, \; i < j, \tag{22}$$

$$= \begin{cases} 1 \text{ if } k = i \text{ or } k = j, \\ 0 \text{ if } k \neq i \text{ and } k \neq j. \end{cases}$$

Now consider $\partial\alpha_1/\partial\theta_b$:

$$\frac{\partial\alpha_1}{\partial\theta_b} = \frac{\partial(\sigma_{12} + \sigma_{13} + \cdots + \sigma_{1K})}{\partial\sigma_{ij}}, \; i < j, \tag{23}$$

$$= \begin{cases} 1 \text{ if } i = 1 \\ 0 \text{ if } i \neq 1. \end{cases}$$

Combining (22) and (23) into (21) we obtain

$$\frac{\partial\gamma_k}{\partial\theta_b} = \frac{\partial\gamma_k}{\partial\sigma_{ij}} = \frac{\alpha_1 P - \alpha_k Q}{\alpha_1^2}, \; i < j, \; b = 1, \cdots, KC, \tag{24}$$

where

$$P = \begin{cases} 1 \text{ if } k = i \text{ or } k = j \\ 0 \text{ otherwise} \end{cases}$$

$$Q = \begin{cases} 1 \text{ if } i = 1 \\ 0 \text{ otherwise} \end{cases}$$

In summary, then, we have

$$\frac{\partial \rho_{F1}}{\partial \theta_b} = \begin{cases} \dfrac{\sigma_X^2 \cdot \dfrac{\partial \sigma_T^2}{\partial \theta_b} - 2\sigma_T^2}{\sigma_X^4}, & b = 1, \cdots, KC \\[2ex] \dfrac{-\sigma_T^2}{\sigma_X^4}, & \text{if } b = KC + 1, \cdots, KV. \end{cases} \tag{25}$$

$$\frac{\partial \sigma_T^2}{\partial \theta_b} = \frac{\left(Y \cdot I \cdot \frac{\partial g}{\partial Y}\right) + \left(\frac{\partial g}{\partial \theta_b} \cdot \alpha_1\right)}{Y^2 \cdot \frac{\partial g}{\partial Y}} \tag{26}$$

where

$$I = \begin{cases} 1 \text{ if } b = 1, \cdots, K\text{-}1 \\ 0 \text{ if } b = K, \cdots, KC \end{cases}$$

$$\frac{\partial g}{\partial Y} = \mp \frac{1}{2\sqrt{.25-Y}} + \sum_{k \neq 1} \frac{\gamma k}{2\sqrt{.25-Y\gamma_k}} \tag{27}$$

$$\frac{\partial g}{\partial \theta_b} = \sum_{k \neq 1} \frac{Y}{2\sqrt{.25-Y\gamma_k}} \cdot \frac{\partial \gamma_k}{\partial \theta_b} \tag{28}$$

$$\frac{\partial \gamma_k}{\partial \theta_b} = \frac{\partial \gamma_k}{\partial \sigma_{ij}} = \frac{\alpha_1 P - \alpha_k Q}{\alpha_1^2}, \quad i < j,\ b = 1, \cdots, KC \tag{29}$$

where

$$P = \begin{cases} 1 \text{ if } k = i \text{ or } k = j \\ 0 \text{ otherwise} \end{cases}$$

$$Q = \begin{cases} 1 \text{ if } i = 1 \\ 0 \text{ otherwise} \end{cases}$$

Finally, by substituting (26) into (25) we obtain

$$\frac{\partial \rho_{F1}}{\partial \theta_b} = \begin{cases} \dfrac{\dfrac{\sigma_X^2 \left[\left(Y \cdot I \cdot \frac{\partial g}{\partial Y}\right) + \left(\frac{\partial g}{\partial \theta_b} \cdot \alpha_1\right)\right]}{Y^2 \cdot \frac{\partial g}{\partial Y}} - 2\sigma_T^2}{\sigma_X^4}, & \text{if } b = 1, \ldots, KC \\ \dfrac{-\sigma_T^2}{\sigma_X^4}, & \text{if } b = KC + 1, \ldots, KV \end{cases} \tag{30}$$

where

$$I = \begin{cases} 1 \text{ if } b = 1, \cdots, K-1 \\ 0 \text{ if } b = K, \cdots, KC, \end{cases}$$

and $\partial g/\partial y$ and $\partial g/\partial \theta_b$ are determined by expressions (27), (28) and (29).

Clearly, the determination of the theoretical variance of r_{F1} involves considerable computation. In the present study computer programs were developed to evaluate the multitude of terms arising for illustrative matrices presented in a later section.

Theoretical Variance of r_{F2}

The population counterparts to the D_j on page 5 will henceforth be represented by β_j . In this notation,

$$\beta_j = \frac{\alpha_j - \sigma_{j\ell}}{\alpha_\ell - \sigma_{j\ell}}, \; j = 1, \cdots, K$$

where α_j is defined as before, namely, the sum of the population covariances in row j.

The population reliability, analogous to equation (4), is

$$\frac{(\Sigma\beta_j)^2 \sum\sum_{i \neq j} \sigma_{ij}}{(\Sigma\beta_j)^2 - \Sigma\beta_j^2} \Bigg/ \sigma_X^2$$

By equation (8) the variance of r_{F2} is

$$\mathrm{Var}(r_{F2}) = \sum_{b=1}^{KV} \sum_{d=1}^{KV} \left[\frac{\partial \rho_{F2}}{\partial \theta_b} \cdot \frac{\partial \rho_{F2}}{\partial \theta_d} \cdot \mathrm{Cov}(X_b, X_d) \right] ,$$

where the θ_i and X_i are as defined as before. Again,

$$\mathrm{Cov}(S_{ij}, S_{mn}) = \frac{\sigma_{im}\sigma_{jn} + \sigma_{in}\sigma_{jm}}{N-1} .$$

What we now need to find is $\partial\rho_{F2}/\partial\theta_b$ for b = 1, $\cdots$, KV. We have

$$\frac{\partial \rho_{F2}}{\partial \theta_b} = \frac{\partial \dfrac{\sigma_T^2}{\sigma_X^2}}{\partial \theta_b} = \frac{\sigma_X^2 \dfrac{\partial \sigma_T^2}{\partial \theta_b} - \sigma_T^2 \dfrac{\partial \sigma_X^2}{\partial \theta_b}}{(\sigma_X^2)^2} , \; b = 1, \cdots, KV.$$

Equations (12) and (13) hold in this situation:

$$\frac{\partial \sigma_X^2}{\partial \theta_b} = \begin{cases} 2 \text{ if } b = 1, \cdots, KC \\ 1 \text{ if } b = KC + 1, \cdots, KV, \text{ and} \end{cases}$$

$$\frac{\partial \sigma_T^2}{\partial \theta_b} = \begin{cases} \dfrac{\partial \sigma_T^2}{\partial \theta_b}, & b = 1, \cdots, KC \\ 0 & , \text{ if } b = KC + 1, \cdots, KV. \end{cases}$$

Equation (14) also holds:

$$\frac{\partial \rho_{F2}}{\partial \theta_b} = \begin{cases} \dfrac{\sigma_X^2 \cdot \dfrac{\partial \sigma_T^2}{\partial \theta_b} - 2\sigma_T^2}{\sigma_X^4}, & \text{if } b = 1, \cdots, KC \\ \dfrac{-\sigma_T^2}{\sigma_X^4} & , \text{if } b = KC + 1, \cdots, KC. \end{cases}$$

We need to find $\sigma_T^2/\partial\theta_b$, for $b = 1, \cdots , KC$, where the true score variance may be written

$$\sigma_T^2 = \frac{\sum\sum_{i \neq j} \sigma_{ij}}{1 - \dfrac{\Sigma \beta_j^2}{(\Sigma \beta_j)^2}} .$$

Then we have

$$\frac{\partial\sigma_T^2}{\partial\theta_b} = \frac{\left(1 - \frac{\Sigma\beta_j^2}{(\Sigma\beta_j)^2}\right) \cdot \frac{\partial\Sigma\Sigma\sigma_{ij}}{\partial\theta_b} - \left(\sum\sum_{i\neq j} \sigma_{ij}\right) \cdot \frac{\partial\left(1 - \frac{\Sigma\beta_j^2}{(\Sigma\beta_j)^2}\right)}{\partial\theta_b}}{\left(1 - \frac{\Sigma\beta_j^2}{(\Sigma\beta_j)^2}\right)^2} . \quad (31)$$

Clearly, we need the two partial derivatives:

$$\frac{\partial \sum\sum_{i\neq j} \sigma_{ij}}{\partial\theta_b}, \text{ and } \frac{\partial\left(1 - \frac{\Sigma\beta_j^2}{(\Sigma\beta_j)^2}\right)}{\partial\theta_b}, \text{ for } b = 1, \cdots, KC.$$

Consider the first:

$$\frac{\partial \sum\sum_{i\neq j} \sigma_{ij}}{\partial\theta_b} = \frac{\partial(2\sigma_{12} + 2\sigma_{13} + \cdots + 2\sigma_{K-1,K})}{\partial\theta_b} \quad (32)$$

$$= 2 .$$

Consider the second:

$$\frac{\partial\left(1 - \frac{\Sigma\beta_j^2}{(\Sigma\beta_j)^2}\right)}{\partial\theta_b} = \frac{\partial\left(\frac{\Sigma\beta_j^2}{(\Sigma\beta_j)^2}\right)}{\partial\theta_b}, \quad b = 1, \cdots, KC. \quad (33)$$

Then

$$\frac{\partial\left(\frac{\Sigma\beta_j^2}{(\Sigma\beta_j)^2}\right)}{\partial\theta_b} = \frac{\Sigma\beta_j^2\,\frac{\partial(\Sigma\beta_j)^2}{\partial\theta_b} - (\Sigma\beta_i)^2\,\frac{\partial\Sigma\beta_j^2}{\partial\theta_b}}{(\Sigma\beta_j)^4} \tag{34}$$

So we need $\partial(\Sigma\beta_j)^2/\partial\theta_b$ and $\partial\Sigma\beta_j^2/\partial\theta_b$, for $b = 1, \cdots, KC$. We then obtain:

$$\frac{\partial(\Sigma\beta_j)^2}{\partial\theta_b} = 2\Sigma\beta_j^2\,\frac{\partial(\Sigma\beta_j)}{\partial\theta_b} = 2\Sigma\beta_j\Sigma\,\frac{\partial\beta_j}{\partial\theta_b} \tag{35}$$

and

$$\frac{\partial\Sigma\beta_j^2}{\partial\theta_b} = \Sigma\,\frac{\partial\beta_j^2}{\partial\theta_b} = \Sigma 2\beta_j\,\frac{\partial\beta_j}{\partial\theta_b} = 2\Sigma\beta_j\,\frac{\partial\beta_j}{\partial\theta_b} .$$ (36)

Now we need only $\partial\beta_j/\partial\theta_b$. We obtain

$$\frac{\partial\beta_j}{\partial\theta_b} = \frac{\partial\left(\frac{\alpha_j-\sigma_{j\ell}}{\alpha_\ell-\sigma_{j\ell}}\right)}{\partial\sigma_{mn}}$$

$$= \frac{(\alpha_\ell-\sigma_{j\ell})\,\frac{\partial(\alpha_j-\sigma_{j\ell})}{\partial\sigma_{mn}} - (\alpha_j-\sigma_{j\ell})\,\frac{\partial(\alpha_\ell-\sigma_{j\ell})}{\partial\sigma_{mn}}}{(\alpha_\ell-\sigma_{j\ell})^2} . \tag{37}$$

We need then $\partial(\alpha_j - \sigma_{j\ell})/\partial\sigma_{mn}$ and $\partial(\alpha_\ell - \sigma_{j\ell})/\partial\sigma_{mn}$. We know $m < n$, since $b = 1, \cdots, KC$. So

$$\frac{\partial(\alpha_j - \sigma_{j\ell})}{\partial\sigma_{mn}} = \begin{cases} 1 \text{ if } (\ell \neq n \text{ and } j = m) \text{ or } (\ell \neq m \text{ and } j = n), \\ 0 \text{ otherwise,} \end{cases} \tag{38}$$

$$\frac{\partial(\alpha_\ell - \sigma_{j\ell})}{\partial\sigma_{mn}} = \begin{cases} 1 \text{ if } (\ell = n \text{ and } j \neq m) \text{ or } (\ell = m \text{ and } j \neq n), \\ 0 \text{ otherwise.} \end{cases} \tag{39}$$

Substituting these results into (37) we obtain

$$\frac{\partial\beta_j}{\partial\theta_b} = \frac{\partial\beta_j}{\partial\sigma_{mn}} = \frac{(\alpha_\ell - \sigma_{j\ell})I - (\alpha_j - \sigma_{j\ell})J}{(\alpha_\ell - \sigma_{j\ell})^2}, \tag{40}$$

where

$$I = \begin{cases} 1 \text{ if } (\ell \neq n \text{ and } j = m) \text{ or } (\ell \neq m \text{ and } j = n) \\ 0 \text{ otherwise} \end{cases}$$

$$J = \begin{cases} 1 \text{ if } (\ell = n \text{ and } j \neq m) \text{ or } (\ell = m \text{ and } j \neq n) \\ 0 \text{ otherwise} \end{cases}$$

Substituting (34), (35) and (36) into (33) we obtain

$$\frac{\partial\left(1 - \frac{\Sigma\beta_j^2}{(\Sigma\beta_j)^2}\right)}{\partial\theta_b} = \frac{2\Sigma\beta_j\Sigma\beta_j^2\,\Sigma\frac{\partial\beta_j}{\partial\theta_b} - (\Sigma\beta_j)^2 2\Sigma\beta_j\,\frac{\partial\beta_j}{\partial\theta_b}}{(\Sigma\beta_j)^4}$$

$$= \frac{2\left(\Sigma\beta_j^2\,\Sigma\frac{\partial\beta_j}{\partial\theta_b} - \Sigma\beta_j\Sigma\beta_j\,\frac{\partial\beta_j}{\partial\theta_b}\right)}{(\Sigma\beta_j)^3}. \tag{41}$$

Summarizing, we have

$$\frac{\partial\rho_{F2}}{\partial\theta_b} = \begin{cases} \dfrac{\sigma_X^2 \cdot \dfrac{2H - \sum\sum_{i\neq j} \sigma_{ij} \dfrac{\partial H}{\partial\theta_b}}{H^2} - 2\sigma_T^2}{\sigma_X^4}, & \text{if } b = 1, \cdots, KC \\ \dfrac{-\sigma_T^2}{\sigma_X^4}, & \text{if } b = KC + 1, \cdots, KV \end{cases} \tag{42}$$

where

$$H = 1 - \frac{\Sigma\beta_j^2}{(\Sigma\beta_j)^2} ,$$

and $\partial H/\partial\theta_b$ is determined from (40) and (41) above.

As in the case of the standard error of r_{F1}, the standard error for r_{F2} involves a large number of terms for even a relatively small value of K. For the illustrative matrices presented later, the evaluation was accomplished via a computer program written for this purpose.

Theoretical Variance of r_K

To apply the delta method to

$$\rho_K = \frac{(\sigma_{12}\sigma_{13} + \sigma_{12}\sigma_{23} + \sigma_{13}\sigma_{23})^2}{(\sigma_{12})(\sigma_{13})(\sigma_{23})\sigma_X^2}$$

equation (8) must be evaluated for the following six parameters and their associated sample estimates:

$$\begin{array}{ll}
\theta_1 = \sigma_{12} & X_1 = S_{12} \\
\theta_2 = \sigma_{13} & X_2 = S_{13} \\
\theta_3 = \sigma_{23} & X_3 = S_{23} \\
\theta_4 = \sigma_{11} & X_4 = S_{11} \\
\sigma_5 = \sigma_{22} & X_5 = S_{22} \\
\theta_6 = \sigma_{33} & X_6 = S_{33} .
\end{array}$$

We again obtain

$$\frac{\partial \rho_K}{\partial \theta_b} = \begin{cases} \dfrac{\sigma_X^2 \dfrac{\partial \sigma_T^2}{\partial \theta_b} - 2\sigma_T^2}{\sigma_X^4}, & \text{if } b = 1,2,3 \\[2ex] \dfrac{-\sigma_T^2}{\sigma_X^4}, & \text{if } b = 4,5,6. \end{cases} \tag{43}$$

This result indicates we must find $\partial\sigma_T^2/\partial\theta_b$ for $b = 1,2,3$. It will be easier to follow the subscripts if σ_T^2 is expressed as a function of the θ_i :

$$\sigma_T^2 = \frac{(\theta_1\theta_2 + \theta_1\theta_3 + \theta_2\theta_3)^2}{\theta_1\theta_2\theta_3}.$$

Letting $A = (\theta_1\theta_2 + \theta_1\theta_3 + \theta_2\theta_3)$ and $B = \theta_1\theta_2\theta_3$, we have

$$\sigma_T^2 = \frac{A^2}{B},$$

$$\frac{\partial\sigma_T^2}{\partial\theta_b} = \frac{B\,\frac{\partial A^2}{\partial\theta_b} - A^2\,\frac{\partial B}{\partial\theta_b}}{B^2}$$

$$= \frac{2AB\,\frac{\partial A}{\partial\theta_b} - A^2\,\frac{\partial B}{\partial\theta_b}}{B^2}$$

$$= \frac{2BA(\Sigma\theta_i - \theta_b) - A^2\,\frac{B}{\theta_b}}{B^2}$$

$$= \frac{2A\left(\frac{\Sigma\theta_i\theta_b - \theta_b^2}{\theta_b}\right) - \frac{A^2}{\theta_b}}{B}$$

$$= \frac{A[2(\Sigma\theta_i\theta_b - \theta_b^2) - A]}{B\theta_b}\ .$$

Now,

$$\text{if } b = 1,\ \frac{\partial\sigma_T^2}{\partial\theta_b} = \frac{A}{B\theta_1}\,(\theta_1\theta_2 + \theta_1\theta_3 - \theta_2\theta_3)$$

$$\text{if } b = 2,\ \frac{\partial\sigma_T^2}{\partial\theta_b} = \frac{A}{B\theta_2}\,(\theta_2\theta_1 + \theta_2\theta_3 - \theta_1\theta_3)$$

$$\text{if } b = 3,\ \frac{\partial\sigma_T^2}{\partial\theta_b} = \frac{A}{B\theta_3}\,(\theta_3\theta_1 + \theta_3\theta_2 - \theta_1\theta_2)\ .$$

Then we have

$$\frac{\partial\sigma_T^2}{\partial\theta_b} = \frac{A}{B\theta_b}\left(A - \frac{2B}{\theta_b}\right)$$

$$= \frac{A^2\theta_b - 2AB}{B\theta_b^2}$$

$$= \frac{\sigma_T^2\theta_b - 2A}{\theta_b^2}, \text{ if } b = 1,2,3.$$

Substituting the foregoing result into (43) we finally obtain

$$\frac{\partial\rho_K}{\partial\theta_b} = \begin{cases} \dfrac{\sigma_X^2\left[\dfrac{\sigma_T^2\theta_b - 2A}{\theta_b^2}\right] - 2\sigma_T^2}{\sigma_X^4}, & b = 1,2,3 \\ \\ \dfrac{-\sigma_T^2}{\sigma_X^4}, & b = 4,5,6 \end{cases} \qquad (44)$$

where $A = \theta_1\theta_2 + \theta_1\theta_3 + \theta_2\theta_3$, and the θ_i are defined on pages 24-25. In evaluating the double summation (8) for Var (r_K), we again use

$$\text{Cov}(S_{ij},S_{mn}) = \frac{\sigma_{im}\sigma_{jn} + \sigma_{in}\sigma_{jm}}{N-1}.$$

We have now derived formulas for the approximate sampling variance of r_{F1} , r_{F2} , and r_K . These expressions involve a large number of terms, since the index of each of the double summations in equation (8) runs from 1 through $K(K+1)/2$, or $K^2(K+1)^2/4$ terms in all. Evaluation of the summation requires the numerical value of the elements of the population variance covariance matrix, quantities which are never known in practice. However, they can be evaluated for any postulated matrix. This has been done for eight illustrative matrices presented in the next section. In the following section we summarize the results of a Monte Carlo study undertaken to corroborate the theoretical approximations.

Illustrative Applications of the Sampling Variance Derivations

The foregoing formulas for the squared standard errors were applied to eight hypothetical measures with the variance covariance matrices presented in Tables 1 and 2. The purpose of these applications was to gain some insight into the comparative stability of the three reliability coefficients. The four matrices in Table 1 presume four-part tests and those in Table 2 eight-part tests. Within each of these values of K two of the hypothetical measures have a population reliability (ρ) of .6 and two others have a reliability of $.83\overline{3}$. Within each pair of matrices associated with a given K and ρ , one matrix represents tau-equivalent parts and the other congeneric parts. (Tau-equivalent parts exhibit homogeneous inter-part covariances; congeneric parts do not, except in the special case of tau-equivalence.) The tau-equivalent parts for these hypothetical measures are not parallel in the classical sense, since the variances of the parts are unequal—a phenomenon which implies unequal error variances.

TABLE I

Population Covariance Matrices: K = 4

	Essential Tau-Equivalent	Congeneric
$\rho = .83\overline{3}$	$\begin{bmatrix} 21.375 & 9.375 & 9.375 & 9.375 \\ 9.375 & 18.375 & 9.375 & 9.375 \\ 9.375 & 9.375 & 15.375 & 9.375 \\ 9.375 & 9.375 & 9.375 & 12.375 \end{bmatrix}$	$\begin{bmatrix} 38 & 18 & 12 & 6 & \lambda_1 = .4 \\ 18 & 21.5 & 9 & 4.5 & \lambda_2 = .3 \\ 12 & 9 & 11 & 3 & \lambda_3 = .2 \\ 6 & 4.5 & 3 & 4.5 & \lambda_4 = .1 \end{bmatrix}$
$\rho = .6$	$\begin{bmatrix} 48.375 & 9.375 & 9.375 & 9.375 \\ 9.375 & 39.375 & 9.375 & 9.375 \\ 9.375 & 9.375 & 29.375 & 9.375 \\ 9.375 & 9.375 & 9.375 & 20.375 \end{bmatrix}$	$\begin{bmatrix} 64 & 18 & 12 & 6 & \lambda_1 = .4 \\ 18 & 43.5 & 9 & 4.5 & \lambda_2 = .3 \\ 12 & 9 & 26 & 3 & \lambda_3 = .2 \\ 6 & 4.5 & 3 & 11.5 & \lambda_4 = .1 \end{bmatrix}$

TABLE 2

Population Covariance Matrices: K = 8

Essential Tau-Equivalent

$$\rho = .83\overline{3} \quad \begin{bmatrix} 10.34375 & & & & & & & \\ & 6.34375 & & & & & & \\ & & 6.34375 & & & & & \\ & & & 7.34375 & & & 2.34375 & \\ & & & & 5.34375 & & & \\ 2.34375 & & & & & 5.34375 & & \\ & & & & & & 4.34375 & \\ & & & & & & & 3.34375 \end{bmatrix} \begin{bmatrix} 14 & 6 & 4.5 & 4.5 & 3 & 3 & 1.5 & 1.5 \\ 6 & 10 & 4.5 & 4.5 & 3 & 3 & 1.5 & 1.5 \\ 4.5 & 4.5 & 7.375 & 3.375 & 2.25 & 2.25 & 1.125 & 1.125 \\ 4.5 & 4.5 & 3.375 & 8.375 & 2.25 & 2.25 & 1.125 & 1.125 \\ 3 & 3 & 2.25 & 2.25 & 4.5 & 1.5 & .75 & .75 \\ 3 & 3 & 2.25 & 2.25 & 1.5 & 4.5 & .75 & .75 \\ 1.5 & 1.5 & 1.125 & 1.125 & .75 & .75 & 2.375 & .375 \\ 1.5 & 1.5 & 1.125 & 1.125 & .75 & .75 & .375 & 1.375 \end{bmatrix} \begin{array}{l} \lambda_1 = .2 \\ \lambda_2 = .2 \\ \lambda_3 = .15 \\ \lambda_4 = .15 \\ \lambda_5 = .1 \\ \lambda_6 = .1 \\ \lambda_7 = .05 \\ \lambda_8 = .05 \end{array}$$

$$\rho = .6 \quad \begin{bmatrix} 28.34375 & & & & & & & \\ & 16.34375 & & & & & & \\ & & 15.34375 & & & & 2.34375 & \\ & & & 18.34375 & & & & \\ & & & & 12.34375 & & & \\ 2.34375 & & & & & 12.34375 & & \\ & & & & & & 9.34375 & \\ & & & & & & & 6.34375 \end{bmatrix} \begin{bmatrix} 32 & 6 & 4.5 & 4.5 & 3 & 3 & 1.5 & 1.5 \\ 6 & 20 & 4.5 & 4.5 & 3 & 3 & 1.5 & 1.5 \\ 4.5 & 4.5 & 16.375 & 3.375 & 2.25 & 2.25 & 1.125 & 1.125 \\ 4.5 & 4.5 & 3.375 & 19.375 & 2.25 & 2.25 & 1.125 & 1.125 \\ 3 & 3 & 2.25 & 2.25 & 11.5 & 1.5 & .75 & .75 \\ 3 & 3 & 2.25 & 2.25 & 1.5 & 11.5 & .75 & .75 \\ 1.5 & 1.5 & 1.125 & 1.125 & .75 & .75 & 7.375 & .375 \\ 1.5 & 1.5 & 1.125 & 1.125 & .75 & .75 & .375 & 4.375 \end{bmatrix} \begin{array}{l} \lambda_1 = .2 \\ \lambda_2 = .2 \\ \lambda_3 = .15 \\ \lambda_4 = .15 \\ \lambda_5 = .1 \\ \lambda_6 = .1 \\ \lambda_7 = .05 \\ \lambda_8 = .05 \end{array}$$

When an illustrative matrix conforms to the tau-equivalent model, each part is of length 1/K by definition. When an illustrative matrix conforms to the congeneric model, the largest part is arbitrarily set equal to four times the length of the smallest part. The lengths of the parts, that is, the values of λ_j , are indicated to the right of each row in the matrix. The true score variance was arbitrarily set equal to 150 for all population matrices.

The approximate standard errors of the three coefficients computed for the eight illustrative matrices are presented in Table 3. When the part-tests are tau-equivalent, the standard error of coefficient alpha (r_α) is also included. It was computed via the delta method rather than from its true sampling distribution, which is $(1-\rho)F_{\nu_1,\nu_2}$, in order to achieve greater comparability to the other standard errors. The divisor of all of the standard errors, $\sqrt{N-1}$, has been omitted from the values in Table 3. Table 4 presents the standard errors evaluated for sample sizes of 50, 100, and 200.

The sample value of r_{F2} and its theoretical standard error will be affected, to a small degree, by the row sum chosen as the divisor in computing D_1 , ⋯, D_K . (The population parameter ρ_{F2} does not depend on this choice.) The results presented here are based on the use of the row one sum as the divisor. However, the variances obtained by the use of other rows are very close to the results obtained here.

The sample value of r_K and the variance of r_K are dependent upon the numerical values of the variances and covariances for the three parts

TABLE 3

Theoretical Standard Errors Times $\sqrt{N-1}$

		$\rho = .6$	$\rho = .83\overline{3}$
Congeneric			
K = 4	r_{F1}	.6985	.2937
	r_{F2}	.6993	.2941
	r_K	.7055	.2961
K = 8	r_{F1}	.6149	.2571
	r_{F2}	.6151	.2571
	r_K	.7055	.2939
Essential Tau-Equivalence			
K = 4	r_{F1}	.6512	.2744
	r_{F2}	.6512	.2744
	r_K	.7765	.3288
	r_α	.6513	.2744
K = 8	r_{F1}	.6052	.2536
	r_{F2}	.6052	.2536
	r_K	.7765	.3288
	r_α	.6053	.2543

TABLE 4

Theoretical Standard Errors for Three Sample Sizes

		$\rho = .6$			$\rho = .8\overline{3}$		
		50	100	200	50	100	200
Congeneric							
K = 4	r_{F1}	.0997	.0702	.0495	.0419	.0295	.0208
	r_{F2}	.0939	.0702	.0495	.0420	.0295	.0208
	r_K	.1007	.0709	.0500	.0423	.0297	.0209
K = 8	r_{F1}	.0878	.0618	.0435	.0367	.0258	.0182
	r_{F2}	.0878	.0618	.0436	.0367	.0258	.0182
	r_K	.1007	.0709	.0500	.0419	.0295	.0208
Essential Tau-Equivalence							
K = 4	r_{F1}	.0930	.0654	.0461	.0392	.0275	.0194
	r_{F2}	.0930	.0654	.0461	.0392	.0275	.0194
	r_K	.1109	.0780	.0550	.0469	.0330	.0233
	r_α	.0930	.0654	.0462	.0392	.0275	.0194
K = 8	r_{F1}	.0864	.0608	.0429	.0362	.0255	.0180
	r_{F2}	.0864	.0608	.0429	.0362	.0255	.0180
	r_K	.1109	.0780	.0550	.0469	.0330	.0233
	r_α	.0865	.0608	.0429	.0363	.0256	.0180

of the measure. Different values can occur from the different ways in which a four-part (or more) test is reduced to three parts. In general, the more nearly equal the parts, the smaller is the standard error. To achieve maximum part-test homogeneity, the four-part tests were reduced to three by combining parts three and four. The eight-part tests were reduced to three by combining parts one with two, three with four, and five through eight.

The results in Table 3 and 4 show that the standard errors of r_{F1} and r_{F2}, each of which maintain the identity of all parts and all inter-part covariances, decrease from K = 4 to K = 8. For example, under a congeneric test with $\rho = .6$ and N = 50, the standard error of r_{F1} is .0997 with K = 4 and .0878 with K = 8. Thus, these coefficients show the same trend as r_α with increasing numbers of parts.

The tables also show that the standard error of r_K is larger than the others. With K = 4, the difference is small. With K = 8, however, the standard error of r_K is thirty per cent greater than that of r_{F1} and r_{F2}. It seems clear that the advantage of r_{F1} and r_{F2} over r_K depends on the difference between the original K and K = 3.

Another result of interest is the comparison of the standard errors of r_{F1} and r_{F2} with each other and with that of r_α. Even under tau-equivalence, the standard errors of r_{F1} and r_{F2} compare favorably to that of r_α. Moreover, the more easily computed congeneric coefficient, r_{F2}, is as stable as r_{F1}. The latter estimate requires an iterative computational procedure which is feasible by hand, but rather tedious.

Monte Carlo Confirmation of the Theoretical Standard Errors

The population matrices in Tables 1 and 2 were used to generate sample matrices through a procedure developed by Odell and Feiveson [1966] and Browne [1968]. Two thousand sample matrices were generated under each of the twenty-four combinations of the following conditions:

Number of part-tests:	$K = 4$ and $K = 8$
Population reliability:	$\rho = .83\overline{3}$ and $.6$
Model:	tau-equivalent and congeneric
Sample size:	$N = 50$, $N = 100$, $N = 200$.

All of the foregoing coefficients were computed for each of 2000 sample matrices under each configuration. In addition, the Mayekawa and Haebara [1980] least squares coefficient for congeneric tests, designated r_{LS} , was computed for comparative purposes for each sample matrix. The empirical standard errors are reported in Table 5. The differences between the Monte Carlo and theoretical standard errors, and the per cent deviation from the theoretical values, are presented in Table 6.

The theoretical approximations by the delta method agree fairly closely with the empirical estimates of the standard errors of r_{F1} , r_{F2} , and r_K . The largest differences occur in instances where negative sample covariances were more likely, that is, with $\rho = .6$ and $N = 50$. (Negative sample covariances must be assumed not to occur for valid application of the delta method.) The same issue probably accounts for the empirical standard errors of r_K being substantially greater than the theoretical

TABLE 5

Standard Errors of Monte Carlo Sampling Distributions

		ρ = .6			ρ = .833̄		
	N =	50	100	200	50	100	200
Congeneric							
K = 4	r_{F1}	.0960	.0719	.0515	.0433	.0307	.0220
	r_{F2}	.0958	.0717	.0515	.0432	.0307	.0220
	r_K	.0972	.0720	.0515	.0435	.0308	.0221
	r_{LS}	.0981	.0730	.0516	.0435	.0307	.0220
K = 8	r_{F1}	.0784	.0613	.0439	.0380	.0270	.0184
	r_{F2}	.0770	.0613	.0438	.0379	.0270	.0184
	r_K	.0952	.0688	.0497	.0434	.0307	.0209
	r_{LS}	.0845	.0640	.0448	.0386	.0273	.0185
Essential Tau-Equivalence							
K = 4	r_{F1}	.0842	.0639	.0464	.0405	.0287	.0203
	r_{F2}	.0830	.0630	.0462	.0404	.0286	.0203
	r_K	.1133	.0919	.0648	.0497	.0349	.0245
	r_α	.0893	.0671	.0476	.0422	.0292	.0206
	r_{LS}	.0845	.0641	.0465	.0405	.0287	.0203
K = 8	r_{F1}	.0754	.0594	.0436	.0375	.0267	.0184
	r_{F2}	.0758	.0581	.0442	.0372	.0266	.0184
	r_K	.1133	.0883	.0646	.0498	.0344	.0243
	r_α	.0775	.0610	.0443	.0387	.0271	.0185
	r_{LS}	.0816	.0607	.0438	.0377	.0267	.0184

TABLE 6

Differences Between Monte Carlo and Theoretical Standard Errors, and Per Cent Deviation from Theoretical Values *

		ρ = .6			ρ = .833		
		N = 50	100	200	50	100	200
Congeneric Measures							
K = 4	r_{F1}	.0038(3.8%)	.0017(2.4%)	.0020(4.0%)	.0013(3.1%)	.0012(4.1%)	.0012(5.8%)
	r_{F2}	-.0041(4.1%)	.0014(2.0%)	.0019(3.8%)	.0012(2.9%)	.0011(3.7%)	.0011(5.3%)
	r_K	.0036(3.5%)	.0011(1.6%)	.0015(3.0%)	.0012(2.8%)	.0010(3.4%)	.0010(4.8%)
K = 8	r_{F1}	-.0095(10.8%)	-.0005(0.8%)	.0003(0.7%)	.0013(3.5%)	.0012(4.7%)	.0002(1.1%)
	r_{F2}	-.0109(12.4%)	-.0005(0.8%)	.0002(0.5%)	.0012(3.3%)	.0012(4.7%)	.0002(1.1%)
	r_K	-.0056(5.6%)	-.0021(3.0%)	.0003(0.6%)	.0014(3.3%)	.0012(4.1%)	.0001(0.5%)
Tau-Equivalent Measures							
K = 4	r_{F1}	-.0088(9.5%)	-.0016(2.4%)	.0002(1.6%)	.0013(3.3%)	.0011(4.0%)	.0008(4.1%)
	r_{F2}	-.0100(10.8%)	-.0025(3.8%)	.0000(0.0%)	.0012(3.1%)	.0010(3.6%)	.0008(4.1%)
	r_K	.0024(2.2%)	.0139(17.8%)	.0098(17.8%)	.0027(5.7%)	.0019(5.7%)	.0012(5.2%)
	r_α	-.0037(4.0%)	.0017(2.6%)	.0014(3.0%)	.0030(7.7%)	.0017(6.2%)	.0012(6.2%)
K = 8	r_{F1}	-.0111(12.8%)	-.0014(2.3%)	.0007(1.6%)	.0013(3.6%)	.0011(4.3%)	.0004(2.2%)
	r_{F2}	-.0107(12.4%)	-.0027(4.4%)	.0013(3.0%)	.0010(2.8%)	.0010(3.9%)	.0004(2.2%)
	r_K	.0024(2.2%)	.0103(13.2%)	.0096(17.5%)	.0028(6.0%)	.0014(4.2%)	.0010(4.3%)
	r_α	-.0090(10.4%)	.0002(0.3%)	.0014(3.3%)	.0024(6.6%)	.0015(5.9%)	.0005(2.8%)

* Computed from data taken to six decimal places

value. With $\rho = .83\overline{3}$ the discrepancy between empirical and theoretical standard errors was less than 6 per cent for all twelve combinations of N and K . Undoubtedly this discrepancy is largely accounted for by the limitations of the delta method.

Conclusions and Recommendations

The theoretical and empirical standard errors lead to the following conclusions:

1) For the congeneric case and K = 4, the standard errors of r_{F1} , r_{F2} , r_K and r_{LS} exhibit no important differences with samples 50 or greater. But with K = 8, the standard error of r_K is substantially larger than those of r_{F1} , r_{F2} and r_{LS} .
2) In almost all cases the standard errors of r_{F1} and r_{F2} are smaller than those of the other coefficients. This finding also holds true under tau-equivalence, in which r_α was included among the comparisons.
3) The coefficients which maintain the identity of the parts and are based on the full set of part-test variances and covariances (r_{F1} , r_{F2} , r_α , r_{LS}) exhibit decreasing standard errors as the number of part-tests increased from four to eight. The standard error of a Kristof coefficient did not depend on the original number of parts before combination.
4) In every comparison the standard errors of r_{F1} and r_{F2} show no material difference. To achieve this approximate equality it may be necessary to use the largest part-test to define the pivot row for r_{F2}.

The bias of all the coefficients was also examined in the Monte Carlo study. None of the coefficients exhibited any substantial bias except coefficient alpha under the congeneric case. The negative bias of r_α in this situation was to be expected.

Which of these coefficients is to be preferred for a measure with congeneric parts depends on the factors of most importance to the researcher. In terms of computational simplicity, r_{F2} and r_K have a distinct advantage over r_{F1} and r_{LS} . (Coefficient r_{LS} is probably impractical without a computer routine for its computation.) In terms of their standard errors r_{F1} , r_{F2} and r_{LS} are preferable to r_K . If uniqueness is considered an advantage--in the sense that for a given set of data arbitrary decisions do not affect the computed reliability coefficient--then r_{F1} and r_{LS} are preferred over r_{F2} and r_K .

As a compromise the authors favor r_{F2} . No disadvantage is associated with this coefficient in terms of standard error. Computation is not difficult once the part-test variance covariance matrix is available. Monte Carlo simulation data suggest that the row with the largest sum of elements serves well as the pivot row. If the researcher is uncertain whether the parts can be assumed to be tau-equivalent or can be regarded to be only congeneric, r_{F2} loses no advantage that coefficient alpha might be thought to have. Coefficient r_{F2} is as adequate as alpha for truly tau-equivalent parts and quite superior to alpha with congeneric parts.

REFERENCES

Browne, Michael W. A comparison of factor analytic techniques. Psychometrika, 1968, 33, 267-334.

Cronbach, L.J. Coefficient alpha and the internal structure of tests. Psychometrika, 1951, 16, 297-334.

Feldt, L.S. Estimation of the reliability of a test divided into two parts of unequal length. Psychometrika, 1975, 40, 557-561.

Fisher, R.A. Moments and product-moments of sampling distributions. Proceedings of the London Mathematical Society, 1928, 30, 199-238.

Gilmer, Jerry S. Reliability estimation of a test with parts of unequal length. Unpublished doctoral dissertation, University of Iowa, 1981.

Gilmer, Jerry S. and Feldt, Leonard S. Reliability estimation for a test with parts of unequal and unknown lengths. Unpublished manuscript, 1981.

Horst, Paul Estimating total test reliability from parts of unequal length. Educational and Psychological Measurement, 1951, 11, 368-371.

Kendall, M.G. and Stuart, A. The Advanced Theory of Statistics: Vol. I, Distribution Theory, (3rd Ed.). London: Charles Griffin and Company, 1969.

Kristof, W. Estimation of reliability and true score variance from a split of a test into three arbitrary parts. Psychometrika, 1974, 39, 491-499.

Mayekawa, S. and Haebara, T. Estimation of the reliability of a test consisting of more than three congeneric parts. Iowa Testing Programs Occasional Papers, Number 28. e University of Iowa, Iowa City, Iowa, 1980.

Novick, M.R. and Lewis, C. Coefficient alpha and the reliability of composite measurements. Psychometrika, 1967, 32, 1-13.

Odell, P.L. and Feiveson, A.H. A numerical procedure to generate a sample covariance matrix. Journal of the American Statistical Association, 1966, 61, 199-203.